



**Conseil économique  
et social**

Distr.  
GÉNÉRALE

CES/2002/7  
5 avril 2002

FRANÇAIS  
Original: ANGLAIS

COMMISSION DE STATISTIQUE et  
COMMISSION ÉCONOMIQUE POUR L'EUROPE

CONFÉRENCE DES STATISTICIENS EUROPÉENS

Cinquantième réunion plénière  
(Paris, 10-12 juin 2002)

**FOURNITURE DE DONNÉES STATISTIQUES  
AUX ORGANISATIONS INTERNATIONALES**

**LE POINT DE VUE D'UN PAYS**

Document présenté par Statistics Norway

**INTRODUCTION**

1. Le système international des statistiques officielles se compose des instituts nationaux de statistique (INS) et de plusieurs organisations internationales. De manière générale, ces composantes se répartissent le travail ainsi: les INS produisent des statistiques nationales tandis que les organisations internationales recueillent et diffusent des statistiques émanant de plusieurs pays et élaborent des agrégats pour les pays membres. Afin d'assurer la comparabilité des données au niveau international, il convient d'harmoniser les concepts, les normes et les méthodes. Par conséquent, l'une des principales responsabilités des organismes internationaux de statistique<sup>1</sup> est d'offrir un cadre pour les travaux d'uniformisation des concepts, des normes et des méthodes. Les organisations en question peuvent également conférer une valeur ajoutée aux chiffres nationaux en procédant systématiquement à des ajustements et une harmonisation destinés à améliorer la comparabilité de ces données au niveau international. Ce processus dépend de la facilité d'accès aux métadonnées pertinentes.

2. Il est de plus en plus important, pour les utilisateurs aussi bien nationaux qu'internationaux, de disposer de statistiques officielles comparables d'un pays à l'autre. Bien souvent, les statistiques nationales ne peuvent être réellement comprises et correctement

interprétées que lorsqu'on les compare à celles d'autres pays, et on a besoin d'un volume toujours croissant d'informations sur les groupements supranationaux à des fins d'analyse des politiques ou de la conjoncture. En réponse à cette demande, la communication d'informations statistiques aux organisations internationales – notamment à Eurostat en tant qu'organisme supranational – a progressivement pris de l'ampleur au cours des 50 dernières années, et, à l'heure actuelle, elle atteint un niveau élevé et couvre la plupart des thèmes d'étude statistiques.

3. Même si le présent document porte essentiellement sur certaines questions concrètes relatives à la communication des données à toute une série d'organisations, il faudrait également prendre en compte les incidences globales de l'adaptation aux besoins internationaux. Parmi les effets positifs, il y a lieu de noter que l'application des normes internationales – considérées, de ce fait, en quelque sorte comme les «meilleures pratiques» – peut simplifier le choix des méthodes et des normes à utiliser au niveau national. Toutefois, il arrive qu'une adaptation stricte des différentes statistiques nationales aux normes internationales rende ces statistiques moins utilisables à des fins nationales. Dans ce cas, les exigences internationales peuvent perturber le processus d'équilibrage qu'un INS se doit de réaliser, compte tenu également de la charge que lui impose la nécessité de communiquer les données et de ses priorités en fonction des ressources disponibles. En outre, les pays étant de dimensions différentes, une norme internationale (fondée sur une sorte de moyenne internationale), par exemple en ce qui concerne la précision, peut compliquer la tâche des petits pays dans la mesure où elle les contraint à élargir la taille de leurs échantillons statistiques.

4. Pour la plupart des INS, répondre aux demandes des organisations internationales représente une charge de travail importante. Ces demandes peuvent accentuer un sentiment de frustration et entraîner des difficultés supplémentaires, en particulier dans les petits INS, si:

- Les normes et définitions ne sont pas bien harmonisées par rapport aux principes nationaux;
- Les normes et définitions ne sont pas bien harmonisées d'une organisation internationale à l'autre;
- Différentes organisations internationales adressent des demandes plus ou moins identiques (entraînant, chacune, des coûts d'expédition);
- Les demandes de données émanant de différentes organisations internationales présentent des divergences arbitraires (non intentionnelles);
- Les organisations internationales ne se servent pas des moyens électroniques de transmission des données;
- Différentes organisations internationales n'utilisent pas la même solution technique pour la transmission des données;
- Les formats de transfert des données ne sont pas bien conçus ni efficaces; et
- Une documentation ou des explications appropriées font défaut et les informations fournies précédemment ne sont pas prises en compte.

5. Par conséquent, il est important de se concentrer sur les mesures et les initiatives susceptibles d'améliorer davantage la qualité et l'efficacité de ce processus de transmission. Le présent document abordera quelques questions générales ainsi que certains cas concrets, et il y a lieu d'espérer que l'on en retirera des idées pour une réflexion plus approfondie et un suivi.

6. Il convient de souligner, en particulier dans le cadre des travaux de la Conférence, que de nombreuses initiatives ont été prises, avec succès, en vue d'améliorer la coordination entre les différentes organisations internationales au cours des 5 à 10 dernières années. Cependant, les tâches des organisations vont s'amplifiant, de nouveaux centres d'intérêt international se font jour, et il faudra donc s'efforcer continuellement d'améliorer la coordination des communications internationales de données. L'objectif à atteindre pourrait être défini comme suit:

La fourniture effective de statistiques aux organisations internationales devrait alourdir aussi peu que possible la tâche des instituts nationaux de statistique. Toutes les demandes devraient être bien formulées, bien coordonnées avec celles d'autres organisations et faire appel au format de transfert le plus efficace. Il faudra à cet effet y inclure impérativement des métadonnées bien coordonnées au niveau international.

## **QUELQUES EXPÉRIENCES NATIONALES**

7. La manière de gérer la coopération entre le système statistique international et le système statistique national dépendra dans une certaine mesure du degré de centralisation de ce dernier. Si le système national est décentralisé, les demandes émanant des organisations internationales peuvent en principe emprunter plusieurs voies de communication, c'est-à-dire qu'elles ne sont pas forcément adressées à l'Institut national de statistique ou ne passent pas nécessairement par son intermédiaire. En Norvège, la loi sur les statistiques confère à Statistics Norway un rôle de coordonnateur de l'activité statistique internationale. Par ailleurs, comme la production des statistiques officielles est assez centralisée en Norvège, nous devrions être en mesure d'avoir une assez bonne vue d'ensemble des statistiques transmises. Toutefois, il y a des domaines pour lesquels l'essentiel de la production de statistiques est effectué au sein des ministères et des directions administratives, Statistics Norway n'ayant actuellement qu'une vue partielle des transmissions de données émanant de ces entités.

8. Il convient de noter que Statistics Norway entretient des relations suivies avec la plupart des organisations internationales, de même qu'avec Eurostat étant donné que la Norvège applique formellement la majeure partie des prescriptions juridiques de la législation européenne en matière de statistique, conformément à l'Accord sur l'EEE. On estime que plus de la moitié des ressources allouées à la production de statistiques au sein de Statistics Norway est consacrée directement ou indirectement à la satisfaction des exigences de la coopération dans le cadre de l'EEE. Ainsi, les incidences de la coopération internationale vont bien au-delà de la quantité relativement faible de ressources attribuable au transfert direct de données (voir ci-après).

9. En 1999/2000, Statistics Norway a effectué une enquête sur la fourniture de statistiques à différentes organisations internationales dans le but de déterminer comment améliorer l'organisation et les processus internes pour répondre aux demandes de ces organisations.



Sujet	Eurostat	OCDE	FMI	ONU	OIT	Comité nordique	Conseil européen	UNESCO	OMS	OMC	CEMT
Statistiques agricoles et forestières et statistiques de la pêche	X										
Statistiques industrielles	X	X		X							
Statistiques de l'énergie	X										
Statistiques de la science et de la technique	X										
Registres des activités industrielles et commerciales											
Statistiques démographiques	X		X	X		X	X				
Recensements de la population et des habitations	X			X			X				
Marché du travail	X	X			X	X					
Statistiques sur l'éducation et la formation	X	X						X			
Statistiques culturelles											
Revenu et dépenses des ménages et bien-être de la population	X		X								
Sécurité sociale											
Santé (y compris les causes de mortalité)	X	X				X			X		
Statistiques différenciées en fonction du sexe											
Indicateurs et cadres sociaux											
Statistiques de l'environnement	X	X									

14. En plus de cette liste d'organisations internationales, il convient de mentionner que la Banque centrale européenne (BCE) recueille de plus en plus de statistiques sur le marché financier auprès des banques centrales nationales et des INS, en partie par l'intermédiaire d'Eurostat. La Banque des règlements internationaux (BRI) collecte elle aussi des statistiques, essentiellement auprès des banques centrales nationales. En outre, il convient de noter que l'Agence internationale de l'énergie (AIE) rassemble périodiquement des statistiques sur l'énergie.

15. Pour ce qui est des ressources directement utilisées pour compiler les données demandées, il a été estimé que les livraisons de données aux organisations mentionnées, dans les domaines énumérés ci-après, exigeaient une assez grande quantité de travail (il se pourrait que certaines communications importantes de données soient omises):

- Statistiques de la santé (OCDE) (470 heures-hommes);
- Statistiques agricoles (Eurostat) (400);
- Statistiques sociales (Eurostat) (250);
- Statistiques de la santé (Comité nordique de statistiques médicales) (NOMESCO) (250);
- Statistiques de l'éducation (questionnaire commun Eurostat/OCDE/UNESCO) (200);
- Statistiques de l'environnement (Eurostat et OCDE) (200);
- Statistiques démographiques (ONU) (160);
- Statistiques démographiques (Conseil européen) (140).

16. Il semble que le processus de transmission fonctionne mieux pour la communication périodique de statistiques à court terme que les livraisons annuelles ou moins fréquentes encore de statistiques structurelles. Toutefois, il convient également de souligner que le volume de ressources à mettre en œuvre pour la communication internationale de données dépend aussi de l'efficacité de l'organisation interne des services nationaux de statistique. À Statistics Norway, jusqu'à présent, ce sont les unités spécialisées qui ont été chargées de la communication internationale des données dans leur domaine de compétence respectif. Elles doivent utiliser leurs système de production et bases de données ou publications ordinaires pour en extraire les données demandées par les organisations internationales. Ainsi, à l'heure actuelle, il n'existe aucune base de données de référence ou de diffusion homogène qui puisse être utilisée efficacement pour la diffusion des données dans tous les domaines thématiques. Cette situation ne nous semble guère idéale et nous avons récemment commencé à mettre au point une nouvelle structure de bases de données de diffusion qui permettra d'améliorer l'efficacité de la communication de données aux organisations internationales.

17. Un élément important pour l'efficacité de la préparation et de la transmission des données est le moyen de transmission utilisé. D'après l'enquête réalisée en 1999/2000, les communications de données de Statistics Norway étaient ventilées comme suit:

	Total	Eurostat	OCDE	Divers
Papier (courrier)	26	8	4	14
Télécopie	2	1	0	1
Disquette	20	15	2	3
Courrier électronique	39	19	13	7
Total	87	43	19	25

18. Même en cas de communication des données par voie électronique (qui devrait être le moyen de transmission préférentiel pour la plupart des envois), le format utilisé peut avoir un impact majeur sur la facilité de préparation de l'extrait (voir l'exemple relatif à l'OCDE). Le format de la communication influera certainement aussi sur la facilité d'emploi des données pour les destinataires – mais ceci n'est pas notre propos. Il convient de noter que le domaine de la transmission des données évolue sans cesse – avec un accroissement rapide du volume de transmission électronique – et les chiffres fournis ci-dessus sont déjà partiellement dépassés.

19. Il y a lieu de signaler que des tableurs ont souvent été utilisés pour la transmission des données. Toutefois, en particulier pour les données structurelles à dimensions multiples, la conception de ces tableurs a des incidences sur la manière dont les données doivent être organisées et le processus de transmission peut s'avérer inefficace. En principe, les tableaux doivent être remplis manuellement, ce qui exige beaucoup de temps. Il arrive également qu'on les copie à partir d'autres fichiers, avec plusieurs risques d'erreurs. Pour ce genre de transmission, il serait plus efficace de disposer d'un format de fichier bien précis et convenu qui puisse être produit à partir de différents types de bases de données locales.

### **DEUX EXEMPLES: LES STATISTIQUES STRUCTURELLES POUR L'INDUSTRIE ET LES SERVICES ET LES INDICATEURS À COURT TERME DE L'INDUSTRIE ET DES SERVICES DE L'OCDE**

20. La communication de données aux fins des enquêtes de l'OCDE concernant les Statistiques structurelles pour l'industrie et les services et les Indicateurs à court terme de l'industrie et des services peuvent illustrer certaines questions ayant trait, à la fois, à la coordination extérieure et interne, à la normalisation des classifications et aux formats des données à fournir.

21. Pour les besoins de l'enquête relative aux Statistiques structurelles de l'industrie et des services, il est nécessaire de fournir un grand nombre de statistiques structurelles portant sur l'industrie et les services ventilées en fonction des rubriques à trois ou quatre chiffres de la CITI rev.3. Ces données sont recueillies à l'aide d'un questionnaire commun mis au point en collaboration avec l'Organisation des Nations Unies pour le développement industriel (ONUDI) et elles constituent la principale source d'information pour la base de données industrielles STAN (analyses structurelles).

22. À notre avis, ce questionnaire soulève plusieurs problèmes:

- L'utilisation de la ventilation de la CITI soulève des difficultés car notre production de statistiques repose sur la NACE;
- La manière dont les données doivent être présentées, c'est-à-dire sous forme de tableur comportant, en tant que rangées, les rubriques de la CITI et neuf années en tête de colonnes, est assez peu commode;
- Les données à communiquer sont assez proches de celles qui doivent être fournies à Eurostat pour les statistiques sur la structure des entreprises, d'où un sentiment que ces deux collectes de données font double emploi.

23. Nous savons que la coordination entre Eurostat et l'OCDE sur ce thème particulier compte parmi les priorités de ces deux organismes et fait l'objet de discussions au sein du Groupe de travail statistique du Comité de l'industrie de l'OCDE.

24. Pour ce qui est des Indicateurs à court terme, des informations sont demandées sur une base trimestrielle et doivent désormais englober également les services.

25. À notre avis, ce processus pâtit à bien des égards des mêmes faiblesses que la collecte des statistiques structurelles pour l'industrie et les services:

- L'utilisation de la CITI rend malaisée l'extraction de certains chiffres;
- Une présentation sous forme de tableur regroupant toutes les classes de la CITI et des données trimestrielles pour plusieurs années alourdit inutilement la charge de travail (un format texte simple serait plus commode, si chaque série comportait son propre code. Il y a lieu de prendre en compte dans ce contexte les enseignements tirés de la transmission de statistiques à court terme à Eurostat au moyen de GESMES/GESMES CB);
- La multiplicité des correspondants et l'absence de coordination interne a posé des problèmes;
- On estime que cette collecte fait double emploi avec la transmission de données semblables à Eurostat.

26. En résumé, on peut dégager de cet exemple les conclusions suivantes:

- Une utilisation concertée des classifications internationales (en l'occurrence la CITI et la NACE) est essentielle;
- L'organisation du format du questionnaire et le réseau de contacts sont d'une grande importance pour améliorer l'efficacité; et
- L'obligation de fournir des données plus ou moins semblables à Eurostat, d'une part, et à l'OCDE, d'autre part, est ressentie comme un double emploi – indépendamment du fait que ces organisations utilisent des classifications différentes.

## **QUELQUES QUESTIONS PRÉCISES**

27. Comme indiqué dans l'introduction, la communication internationale des données constitue une partie importante de l'activité des INS, et ces derniers mettent tout en œuvre pour établir des concepts, normes et méthodes harmonisés au niveau international et pour adapter les produits statistiques nationaux aux exigences internationales. Il serait regrettable que le processus de transmission et de diffusion de statistiques comparables au niveau international se heurte à des blocages et entrave, en fait, l'accès à des données de cette nature. L'un des aspects de ce processus est la transmission par les systèmes nationaux de statistique des données qu'ils communiquent aux organisations internationales. Un autre aspect tout aussi important est la possibilité tant pour les INS que pour d'autres utilisateurs nationaux et internationaux d'avoir

accès aux données stockées dans les organisations internationales. Les questions abordées ci-après comptent parmi celles qui, d'après notre expérience, méritent de retenir l'attention si l'on veut améliorer l'efficacité de la communication des données aux organisations internationales.

#### Coordination des classifications/normes

28. Il va de soi que si les organisations internationales n'utilisent pas les mêmes classifications, normes et listes de codes, ce manque d'harmonisation se traduira par un surcroît de travail considérable pour les INS. Par conséquent, les organisations devront trancher d'un commun accord les questions irrésolues – où prendre en charge une plus grande part du travail d'harmonisation des données basées sur des classifications différentes. En reprenant l'exemple ci-dessus, l'OCDE pourrait être tenue d'accepter des données classées en fonction de la NACE et de les mettre en concordance avec la CITI – si elle continue d'utiliser cette dernière.

#### Métadonnées communes et claires

29. Il est fondamental que les données à fournir soient clairement définies pour que le processus de production se fasse efficacement et correctement. Le contenu des métadonnées devrait être précisé d'un commun accord et se conformer à certaines normes minimums. Les descriptions des données à fournir devraient autant que possible être attachées aux questionnaires ou y être incorporées, lorsqu'il s'agit de demandes ponctuelles, tandis que dans le cas des transmissions périodiques, il suffit de s'entendre une fois pour toutes sur le contenu des métadonnées, qui pourront être complétées ultérieurement par des révisions. Il en résulte que des longs documents Word comportant des métadonnées et des «questionnaires» sous forme de tableurs ne sont pas considérés comme une solution idéale.

#### De la transmission passive à la transmission active

30. Bon nombre de transmissions de données reposent sur l'hypothèse que les données devraient être transmises avant une date fixe, comme indiqué, par exemple, dans un acte juridique d'Eurostat ou comme convenu de manière informelle. Pour les transmissions périodiques, il faudrait généralement se mettre d'accord sur une date de transmission normale, ce qui simplifierait la programmation des livraisons de données. Toute modification de la couverture des données devrait être notifiée dans un certain délai.

#### Structure des demandes/transmissions

31. Les demandes de données devraient être subdivisées en entités assez logiques – pour faciliter la tâche des différentes organisations. Par conséquent, les demandes d'ampleur importante, pour toutes les statistiques conjoncturelles ou structurelles, ne sont pas pratiques et devraient être fractionnées en unités plus faciles à manipuler. Comme il est souvent difficile pour les INS centralisés de traiter ce genre de demandes complexes, on peut s'attendre à ce qu'il soit encore plus difficile d'y répondre dans des systèmes décentralisés. Le fractionnement des demandes en unités logiques permettra aussi de fournir des descriptions de métadonnées communes et d'employer un format d'échange commun.

### Formats de transfert

32. Les organisations devraient s'entendre sur un format de transfert commun. Comme GESMES/GESMES CB est désormais utilisable pour les transmissions de données d'Eurostat, et appliqué pour la majeure partie des communications de statistiques conjoncturelles adressées par la Norvège à ce dernier, on pourrait sérieusement envisager de retenir cette solution (pour plus de détails, voir la liste des références et l'annexe). En tout état de cause, il faudrait s'abstenir d'utiliser des grands tableurs complexes. Les travaux actuellement menés au titre de l'initiative concernant l'échange de données et métadonnées statistiques (SDMX), lancée par le FMI avec la participation de la BRI, de la BCE, d'Eurostat, de l'OCDE et de l'ONU, donnent à penser que l'on s'engage fermement sur cette voie (voir également l'annexe).

### Traitement des données nouvelles et révisions

33. Il faudrait se borner à demander des mises à jour ou révisions de données déjà fournies. Les données transmises antérieurement pourraient être incorporées dans l'information de retour communiquée dans la demande de nouvelles données (l'exemple précité de l'OCDE).

### Coordination des ensembles de données communs

34. Lorsqu'il y a chevauchement entre les demandes d'ensembles de données de différentes organisations, les informations ne devraient en principe être transmises qu'à une seule d'entre elles – si l'on n'a pas convenu d'un format commun de transmission permettant des communications parallèles. Dans le cas des domaines dont s'occupe Eurostat, on pourrait s'attendre à ce que tous les membres du Système statistique européen (Espace économique européen plus – progressivement – les pays candidats à l'adhésion) se bornent à transmettre leurs données à Eurostat pour les domaines en question.

### Organisation nationale et interne

35. Afin d'améliorer l'efficacité générale de l'échange de données, il importe aussi de passer en revue l'organisation nationale du système de production de statistiques ainsi que l'organisation interne des INS. Il peut s'avérer crucial sur le plan stratégique de mettre sur pied des bases de données de référence et de diffusion qui s'appuient sur des métadonnées normalisées et des instruments types d'échange de données. Les bases de données de cette nature pourraient être soit centralisées à l'intérieur d'un pays, soit fondées sur une formule quelconque de réseau commun dans un système décentralisé. Il importe en outre de mettre en place un réseau de correspondants responsables des différents domaines traités, et d'assurer une bonne coordination entre les différentes disciplines.

### Accès aux stocks de statistiques des organisations internationales

36. Pour créer de manière générale davantage d'incitation à fournir des données aux organisations internationales et pour renforcer la qualité des statistiques tant internationales que nationales, il est essentiel que les INS et d'autres fournisseurs de données aient facilement accès à ces informations. Tout d'abord, les fournisseurs de données devraient y avoir accès sans restriction de manière à pouvoir vérifier les chiffres. Mais il importe également que ces données soient diffusées sous une forme qui permette de les utiliser pour les publications nationales et de les y intégrer, et qu'il n'y ait aucune restriction formelle à une telle intégration. Les comparaisons internationales apportent une valeur ajoutée considérable aux statistiques

nationales, et devraient faire partie intégrante des accords d'échange normaux. Cette condition nécessitera la mise au point d'une politique de tarification commune entre les INS et les organisations internationales. Si les INS envoient des données gratuitement aux organisations internationales, il sera difficile d'admettre que celles-ci fassent payer en retour les ensembles de statistiques communs qu'elles élaboreront à partir de ces données. Dans le cas d'Eurostat, des formules d'échange de cette nature faisant appel à la base de données New Cronos peuvent être établies dans le contexte de l'accord sur les Datashop d'Eurostat, tandis qu'en ce qui concerne l'OCDE, l'accès aux données statistiques est assez restreint ou trop coûteux. L'accès gratuit fondé sur la base de données OLIS est jugé peu commode et trop rigide, notamment parce que ce système s'accompagne d'un degré élevé de sécurité le rendant plus difficile à utiliser que l'Internet. En outre, le coût de l'accès aux données statistiques provenant de la base de données Source de l'OCDE est considéré comme trop élevé.

## CONCLUSIONS ET RECOMMANDATIONS

37. L'une des conditions d'une transmission plus efficace au niveau national semble être la mise sur pied et l'exploitation de bases de données de référence regroupant les données qui doivent être transmises aux organisations internationales, y compris des métadonnées bien structurées et définies d'un commun accord, et permettant de transmettre des données sous une forme standard, s'accompagnant de métadonnées.

38. Les organisations internationales pourraient poursuivre l'examen du sujet traité dans le présent document en s'attachant notamment à quelques points importants, à savoir:

- Veiller à ce que les questions encore en suspens relatives à l'harmonisation des définitions et des classifications soient étudiées et résolues. Le règlement de ces questions sera fondamental pour la mise au point de solutions communes en matière de transmission de données et métadonnées;
- Poursuivre les travaux ayant pour objet de convenir de voies de transmission communes et d'une division du travail. Un pays ne devrait être appelé à fournir un même ensemble de données qu'à une seule organisation;
- Mettre en place des méthodes et instruments sûrs et opérationnels pour la transmission d'informations sensibles. Il faudrait prendre en considération dans ce contexte les enseignements tirés de la solution STADIUM/STATEL appuyant GESMES/CB; et
- Offrir tant aux INS qu'aux autres sources de données un accès convivial aux statistiques disponibles dans les organisations internationales sous une forme permettant des comparaisons et l'exploitation de ces données dans les programmes de diffusion nationaux.

## NOTE

<sup>1</sup> Par «organismes internationaux de statistique», on entend en l'occurrence toutes les organisations internationales qui sont chargées d'un programme ou d'une activité régulière spécifique dans le domaine de la statistique, ou qui demandent des informations statistiques.

## **Bibliographie**

Andravisaneas, C. & Yancis, L.: Exchanging statistical metadata in the European system of central banks (Échange de métadonnées statistiques dans le contexte du système européen de banques centrales). Document présenté lors de l'atelier sur les métadonnées de 2001. Eurostat

Brünger, Heinrich et Svein Longva (1998): Principes internationaux régissant la statistique officielle au niveau national: sont-ils aussi applicables aux travaux statistiques des organisations internationales? Conférence des statisticiens européens, quarante-sixième réunion plénière, 1998 (CES/1998/22)

Coordination de la collecte de données statistiques auprès des pays. Commission de statistique, trente-troisième session, 5-8 mars 2002 (E/CN.3/2002/30)

Ward, D & Pellegrino, M.: Developing a common understanding of standard metadata components: A statistical glossary (Élaboration d'une conception commune des composantes types des métadonnées). Document présenté lors de la réunion de travail commune CEE-ONU/Eurostat sur les métadonnées statistiques, 6-8 mars 2002

## **ANNEXE: EXPLICATIONS COMPLÉMENTAIRES SUR CERTAINES QUESTIONS**

### **GESMES**

1. GESMES est l'acronyme du Generic Statistical Message (message statistique générique), qui a été mis au point par un groupe d'organismes de statistique européens collaborant au sein de l'organe international de normalisation EDIFACT. GESMES présente toutes les caractéristiques requises pour permettre l'échange de tableaux multidimensionnels et de séries chronologiques, y compris de métadonnées (et de notes de bas de page). Il peut également servir à la transmission de données mises en tableaux simples. L'avantage qu'offre GESMES, par rapport à un format de données exclusif, tient au fait qu'il s'agit d'une norme convenue au niveau international, à la fois ouverte et entièrement fonctionnelle. GESMES n'est pas lié au format et aux contraintes d'une application donnée. En particulier, il permet l'échange des catégories d'information suivantes:

- Métadonnées;
- Tableaux multidimensionnels;
- Séries chronologiques; et
- Données administratives.

2. Sous sa forme générique, il permet aussi à l'institution expéditrice d'utiliser la structure d'ensemble de données qu'elle préfère. En 1995, GESMES a été accepté en tant que message de statut 1 d'EDIFACT/ONU et a été publié pour la première fois dans le répertoire D95A de l'ONU.

### **Conditions à satisfaire pour utiliser le format simple de GESMES**

3. On peut utiliser le format simple de GESMES pour expédier n'importe quel type de données ou tableaux multidimensionnels, mais les informations envoyées doivent satisfaire aux conditions suivantes:

- Les concepts statistiques auxquels elles se réfèrent doivent être connus à la fois de l'expéditeur et du destinataire;
- L'ordre des dimensions et des cellules doit être fixé et connu par avance; et
- Le système de codification utilisé pour les valeurs statistiques doit être connu à la fois de l'expéditeur et du destinataire.

4. Pour un complément d'information sur GESMES, voir l'adresse Internet suivante:  
<http://forum.europa.eu.int/Public/irc/dsis/eeg6>.

### **GESMES/CB**

5. GESMES/CB est un message conçu pour permettre l'échange d'informations statistiques entre des organisations, indépendamment de la plate-forme utilisée. Le message met en œuvre

un **modèle d'échange de données de séries chronologiques** (modèle de données GESMES/CB) au moyen duquel on peut échanger des séries chronologiques multidimensionnelles et toute une gamme de métadonnées connexes. Il fait appel à un profil GESMES approprié et à la syntaxe EDIFACT. Tandis que GESMES constitue un modèle générique de données statistiques, qui offre suffisamment de souplesse pour pouvoir décrire de manière syntaxique presque tous les modèles de données statistiques possibles, GESMES/CB présente une syntaxe fixe. Ainsi, les institutions partenaires peuvent concevoir et mettre au point les applications nécessaires pour «lire» et «écrire» des messages GESMES/CB, en évitant les fichiers intermédiaires et les traducteurs spéciaux; la conception des applications lecture/écriture est encore simplifiée par l'élimination du caractère générique superflu lorsqu'il s'agit d'échanger des séries chronologiques. Du fait de la syntaxe fixe, dans la plupart des cas, les règles appliquées dans le cas de GESMES/CB sont plus rigoureuses et restrictives que celles qui régissent le GESMES générique. Cependant, la conception actuelle de GESMES/CB offre la possibilité de procéder à des perfectionnements ultérieurs et une généralisation progressive, si nécessaire, à condition que les parties intéressées en conviennent.

6. GESMES/CB présente plusieurs caractéristiques:

- Adaptation aisée à n'importe quel domaine économique et couverture souple de tous les types de données économiques – statistiques;
- Techniques de représentation modernes: clefs multidimensionnelles conceptuellement «nettes»;
- Efficacité: il évite la répétition inutile d'informations;
- Clefs de séries sans restriction de longueur;
- Introduction aisée d'attributs nouveaux;
- Attributs à divers niveaux (observation, série chronologique et niveau «supérieur»);
- Échange de métadonnées «riches», s'appuyant sur un mode d'organisation souple et efficace;
- Il se prête à l'échange réciproque de séries chronologiques: c'est-à-dire à la communication des données et à leur diffusion;
- Il permet la diffusion *sans support papier* de toute une base de données statistiques: les données, métadonnées, définitions, structures des clefs et listes de codes peuvent être transmises par voie électronique à partir du centre vers les autres institutions;
- Ce message se prête à une mise en application par étape: on peut commencer par les parties essentielles (servant à la transmission des données administratives et numériques) puis employer l'ensemble du message à un stade ultérieur;
- Facilité d'application: il ne devrait pas être nécessaire d'acheter un logiciel spécial à cette fin;

- Conformité aux normes internationales (EDIFACT);
- Large groupe d'utilisateurs: banques centrales nationales du monde entier, Banque des règlements internationaux (BRI), Fonds monétaire international (FMI), plusieurs instituts nationaux de statistique (INS), Eurostat-BoP (statistiques de la balance des paiements), Banque centrale européenne (BCE); et
- Compatibilité avec les objectifs à long terme des institutions partenaires (Système européen de banques centrales, BRI et Eurostat) en matière d'interopérabilité.

7. Pour un complément d'information sur GESMES/CB, voir l'adresse Internet suivante: [www.ecb.int](http://www.ecb.int) (cliquez sur «Statistics», puis sur «GESMES/CB»).

### **LE MESSAGE GÉNÉRIQUE GESMES ET LES NORMES EDIFACT**

8. Le processus de normalisation entrepris dans le cadre d'EDIFACT/ONU a pour but de doter les objets des flux d'échange de données d'une structure dont les éléments peuvent être compris et traités au moyen de programmes informatiques sans intervention humaine. Les premiers messages ont été créés pour échanger des données commerciales et concernaient des catégories d'objets assez simples et statiques comme des factures, des commandes, etc. Par la suite, des secteurs nouveaux, comme celui des statistiques, comportant des structures de données plus complexes et variées, sont venus s'ajouter au groupe des utilisateurs et il est apparu clairement que l'on devait donner un caractère plus générique à certains types de messages pour répondre aux besoins des utilisateurs des secteurs en question sans provoquer pour autant une explosion du nombre de messages.

9. L'Office statistique de l'Union européenne, Eurostat, qui a dirigé la mise au point des messages statistiques EDIFACT/ONU utilise GESMES pour ses échanges de données avec les États appartenant à l'Espace économique européen (EEE) et encourage d'autres organisations internationales et d'autres secteurs à se servir de ces messages.

10. Pour un complément d'information sur l'EDIFACT, voir l'adresse Internet suivante: [www.unece.org/trade/untdid/](http://www.unece.org/trade/untdid/).

### **LE PROJET SDMX (ÉCHANGE DE DONNÉES ET MÉTADONNÉES STATISTIQUES)**

11. La BRI, la BCE, Eurostat, le FMI, l'OCDE et l'ONU ont uni leurs efforts pour rechercher les meilleures pratiques en matière d'information statistique qui leur permettraient d'augmenter l'efficacité des processus d'échange et de partage de données et de métadonnées dans le cadre actuel de leurs activités collectives.

12. L'objectif de cette collaboration est d'étudier des normes électroniques communes et les activités de normalisation en cours qui permettraient d'obtenir un gain d'efficacité et d'éliminer les doubles emplois dans les travaux des organisations concernées elles-mêmes et, éventuellement, dans les travaux d'autres entités en matière d'information statistique.

13. À cet effet, il s'agit de tirer parti des éléments suivants (tant ceux qui existent déjà que ceux qui sont en train d'apparaître):

- Les protocoles d'échanges, tels que GESMES/CB appliqué par les banques centrales pour l'échange de séries chronologiques;
- Les formats de diffusion, comme celui qui est implicite dans le Tableau d'affichage des normes de diffusion des données du FMI; et
- Les normes électroniques, comme le langage XML (eXtensible Markup Language).

14. Pour un complément d'information sur le projet, voir l'adresse Internet suivante: [www.sdmx.org](http://www.sdmx.org). On trouve une description de la version préliminaire du glossaire des termes statistiques SDMX dans Ward, D & Pellegrino, M (2002).

-----